

Cloud Computing

Juan E. Vargas

Principal Research Manager

Microsoft Research (eXtreme Computing Group)

Microsoft

Outline

- Innovation and Cloud Computing
- Windows Azure **Platform** Overview
- MSR/XCG Academic Research Engagements

100 Years of Innovation

- **Health:** public sanitation, aspirin, antibiotics, vaccines, lasers, organ transplants, medical imaging, genome, genomics, epigenetics, cancer genomics (TCGA consortium)....
- **Energy:** electricity, electric grids, nuclear, ... green. . .
- **Transportation:** roads, airplanes, helicopters, ... space exploration...
- **Communications:** radio, TV, phone, mobile phones, . .
- **Electronics:** transistor, computers , internet, WWW, ..
- **Computing has been at the center of innovation during the last 50 years...**
- **Successful technologies are those that become “invisible”**

Pace of Innovation

- It took about **55 years** to spread the use of automobiles to $\frac{1}{4}$ of the US population...
- ... **35 years** for the telephone ...
- ... **22 years** for the radio ...
- ... **16 years** for the PC ...
- ... **13 years** for the cell phone ...
- ... **7 years** for the Internet...
- ... **3 years** for “cloud computing” to take center stage

Cloud Computing

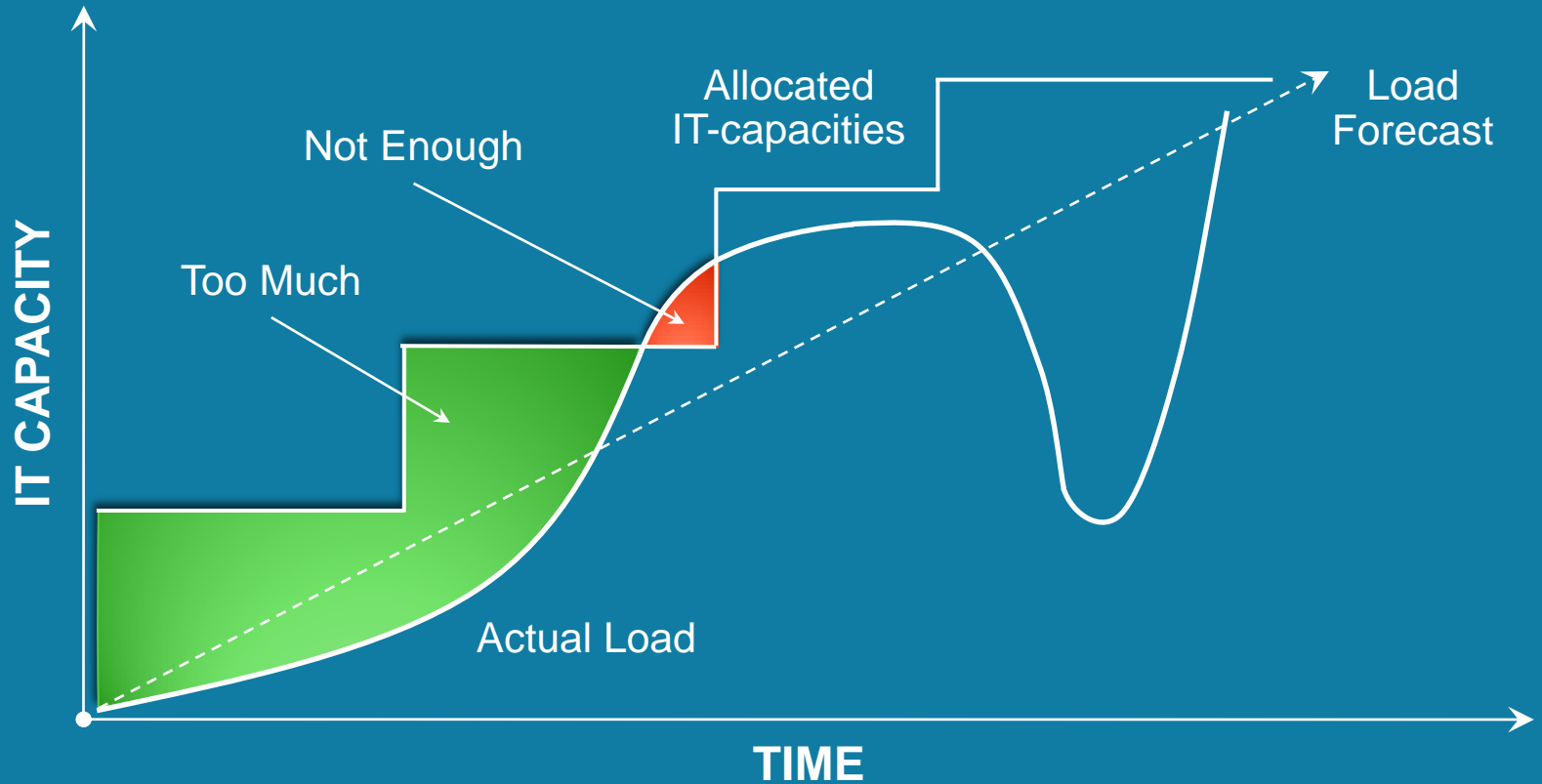
Large scale and on-demand computing via the internet, serving a variety of loads, connecting to a variety of devices and end points.



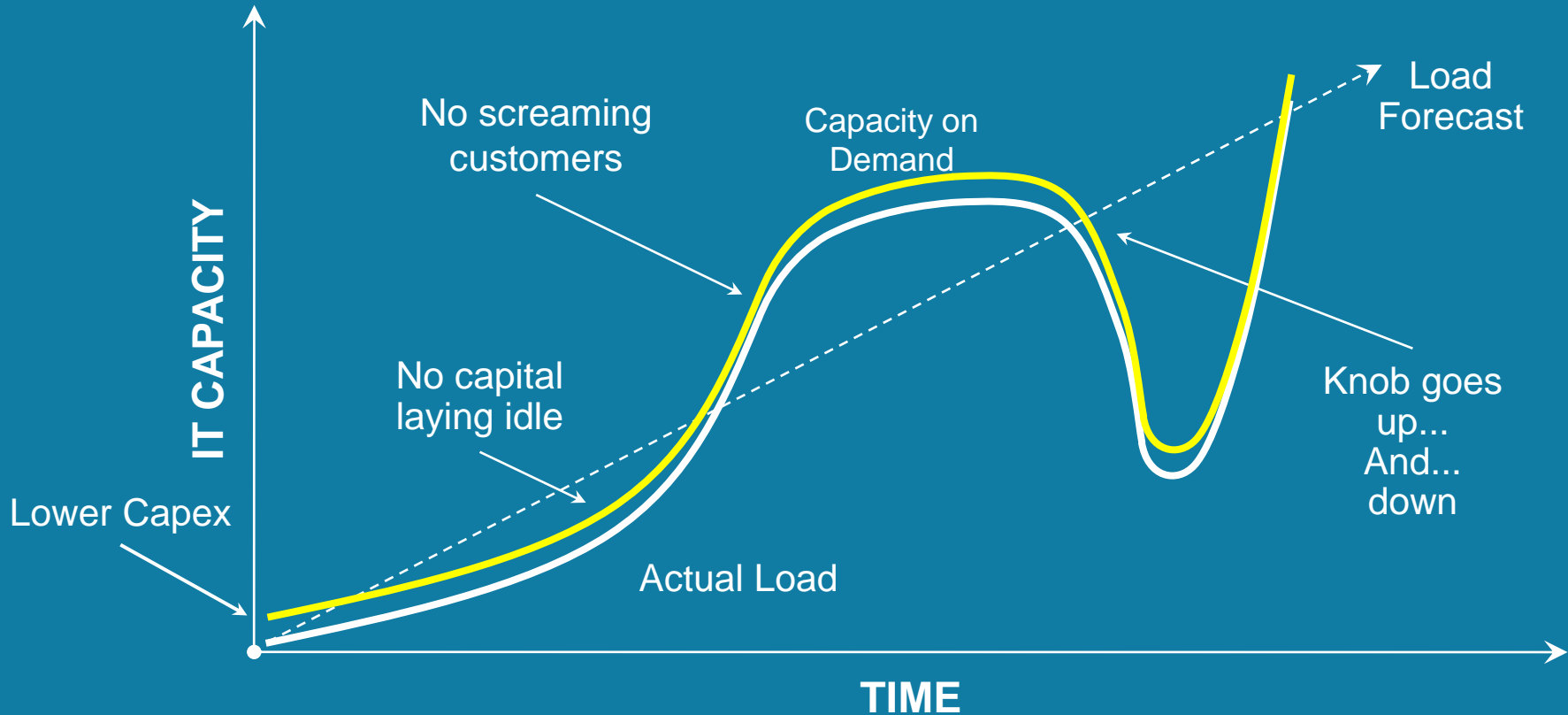
- A **model** of **computation** and **storage** based on **elastic**, (pay as you go) access to vast remote data center capabilities
- An **infrastructure** that provides a framework to manage scalable, reliable, on-demand access to applications
- The "**invisible**" backend to many applications, including mobile
- Historical roots in today's Internet apps
 - Search, email, social networks
 - Files storage (Live Mesh, Mobile Me, Flickr,...)
- A way for **organizations** and for **individuals** to build scalable web presence without making huge investments in IT infrastructure



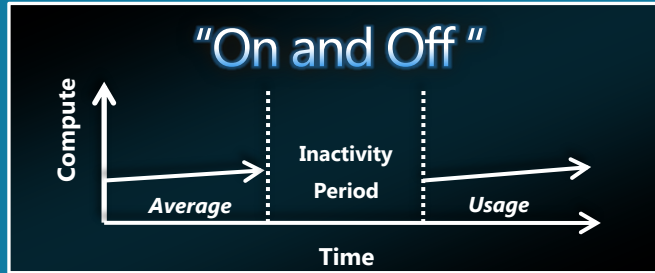
The IT Dilemma



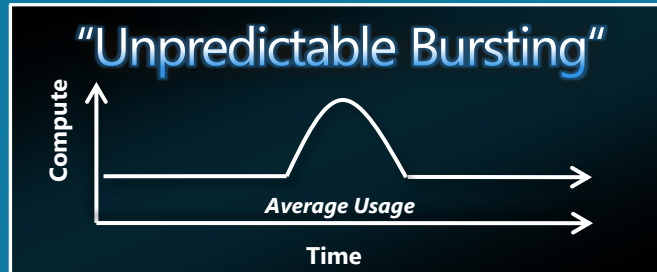
Adapting Capacity to Demand



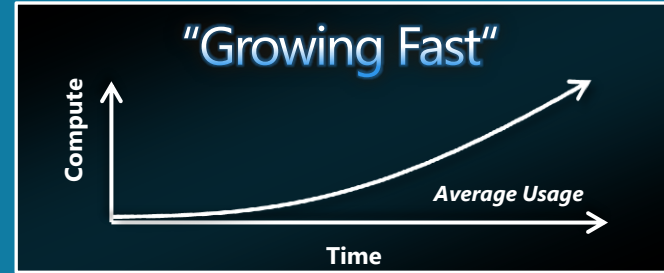
Cloud Computing Patterns



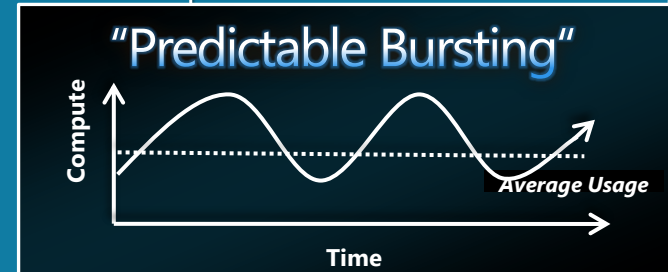
- On & off workloads (e.g. batch job)
- Over provisioned capacity is wasted
- Time to market can be cumbersome



- Unexpected/unplanned peak in demand
- Sudden spike impacts performance
- Can't over provision for extreme cases



- Successful services need to grow/scale
- Keeping up w/ growth is big IT challenge
- Cannot provision hardware fast enough



- Services with micro seasonality trends
- Peaks due to periodic increased demand
- IT complexity and wasted capacity

Cloud Services



"IaaS"

Infrastructure-as-a-Service
host



"PaaS"

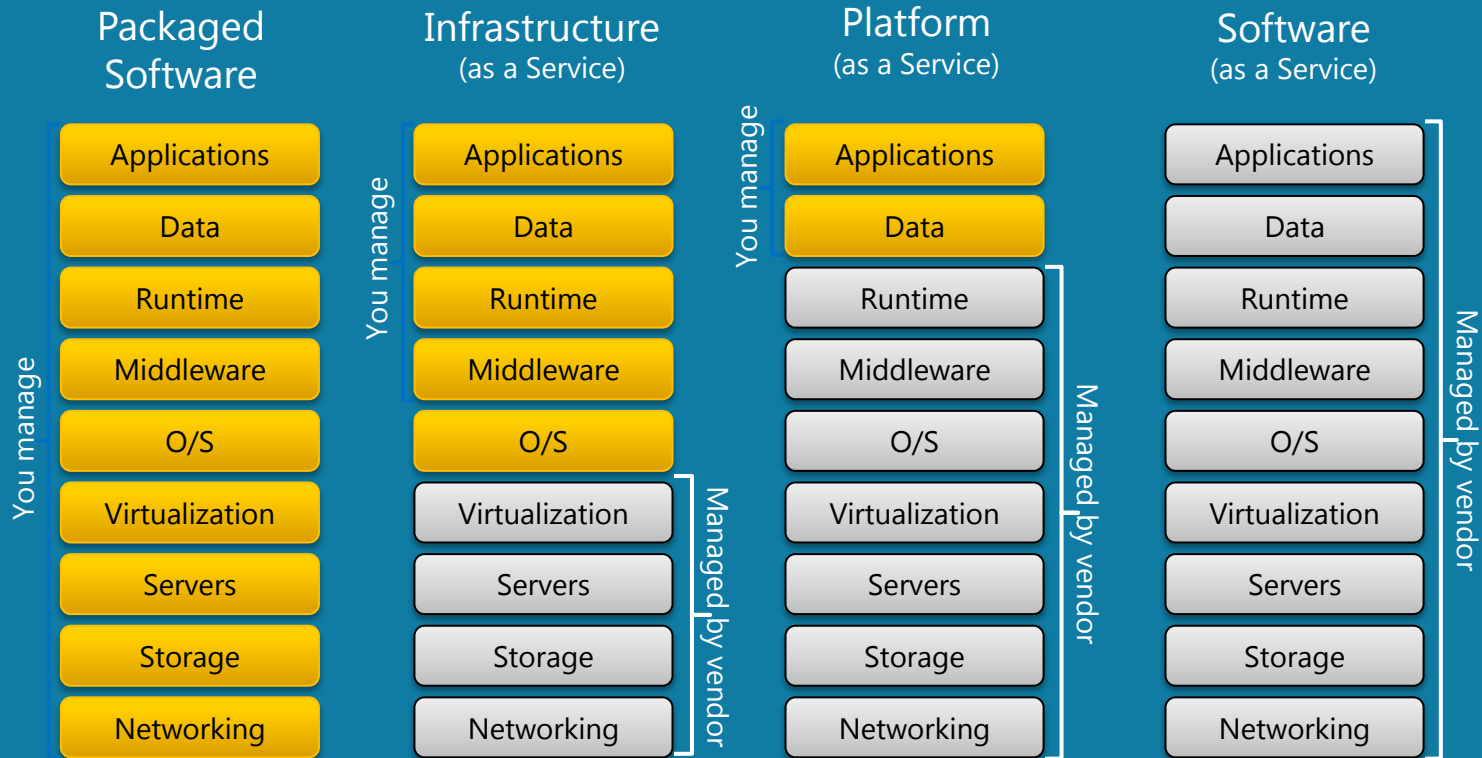
Platform-as-a-Service
build



"SaaS"

Software-as-a-Service
consume

Cloud Services



Today Clouds are Built on Huge Data Centers

- Range in size from “edge” facilities to megascale.
- Economies of scale
 - Approximate costs for a small size center (100-1000 servers) and a larger, 100K or more server center.

| Technology | Cost in small-sized Data Center | Cost in Large Data Center | Ratio |
|----------------|---------------------------------|-----------------------------|-------|
| Network | \$95 per Mbps/month | \$13 per Mbps/month | 7.1 |
| Storage | \$2.20 per GB/month | \$0.40 per GB/month | 5.7 |
| Administration | ~140 servers/Administrator | >1000 Servers/Administrator | 7.1 |



Modern data centers are
about 11.5 times
the size of a football field

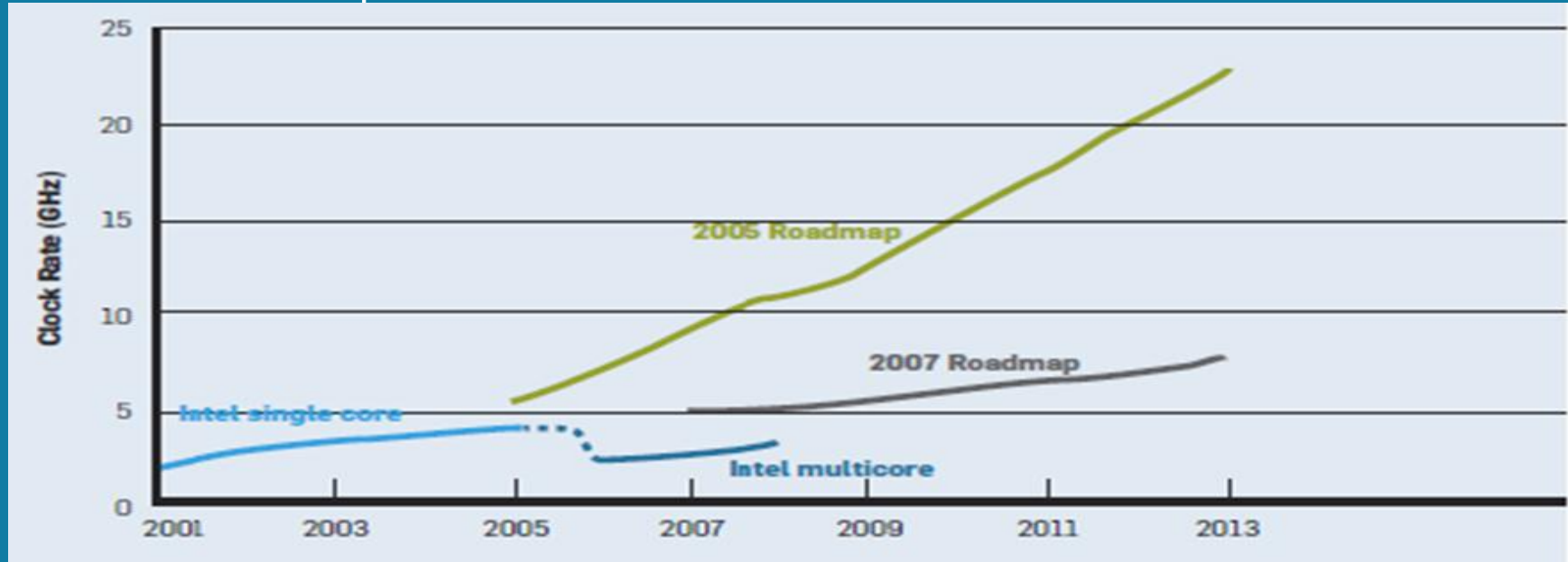
Advances in DC Deployment

- Conquering complexity
 - Building racks of servers & complex cooling systems all separately is not efficient.
 - Package and deploy into container units:



Once-in-a-life-time opportunity

- Increasing performance lead to CPUs that were fast but inefficient.
- CPU clock rates at the (power) limit of what a single chip can dissipate.
- Parallel computing {multi,many}-core is the most promising approach to increase performance

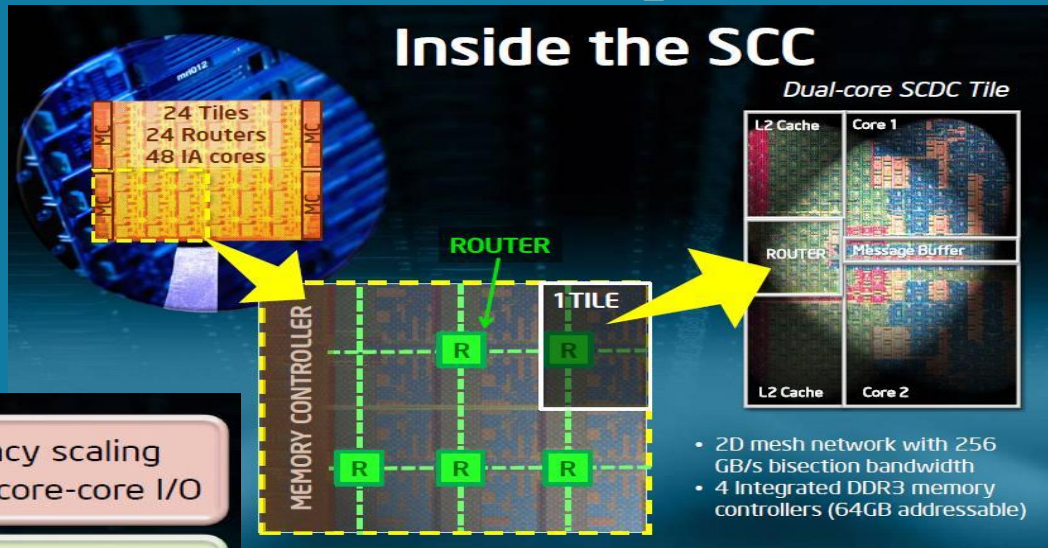


Once-in-a-life-time opportunity

- Parallel computing is a fundamental paradigm shift that breaks the 50-year tradition of software development. Research in this area could have deep implications for applications, programming models, architectures, and OS
- New **applications** that exploit the computational power of parallelism and concurrency at the **client**
- Develop new types of programming models, languages, and tools that can be used to build those applications
- Identify the architectures that will be able to support the new generations of programming models and applications
- Understand how the increased numbers of (possibly heterogeneous) processing units could be combined into a single system

Single-Chip Cloud Computer

December 2009



Energy Efficiency

- Dynamic voltage/frequency scaling
- 1/3 power reduction for core-core I/O

Design Complexity

- Array of small IA-based tiles could lead to more agile, flexible designs

Programming Models

- Message-passing, shared virtual memory, map-reduce, and actors

Application Development

- Working with Microsoft & others for academic, industry innovation

Windows Azure Platform



- Scalable compute and storage
- Automated service management
- Familiar tools, technologies, languages



- Relational storage for the cloud
- Consistent development model
- Automated database management



- Connect existing apps to the cloud
- Connect through network boundaries
- Easily control authorization to apps



Windows Azure Platform



Compute

Storage

Management

Connect



Database

Reporting

Data Sync



Access Control

Service Bus

Caching

demo

Hello Azure

Demo: Hello Azure

What you saw...

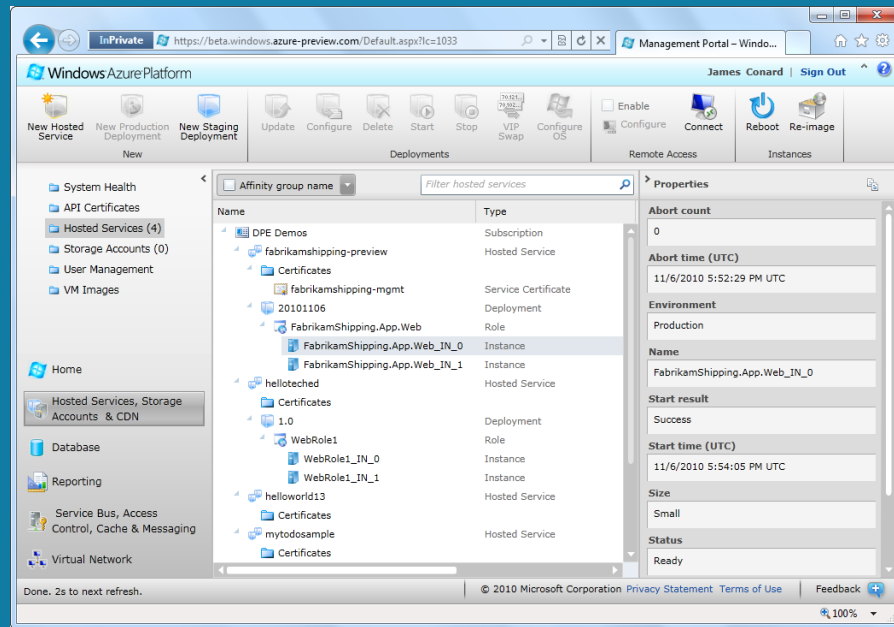
- Simple web app
- Visual Studio 2010
- Roles & instances are models in XML config
- Local F5 debugging
- Deployed to the cloud
- Management Portal
- Switched from staging to production

What Windows Azure provided

- Environment to design, test, and run the app
- Machines, rack space, switches, connectivity
- Automated deployment & configuration
- Isolation, redundancy, load balancing
- Abstraction & Flexibility

Windows Azure Platform Portal

- User friendly portal for Windows Azure Platform services
- More visibility and control
- Supports multiple service administrators (co-admins)
- Enable and connect to Windows Azure instances using Remote Desktop



Tools needed

- VS2010 with SP1 (KB983509)
- <http://www.microsoft.com/windowsazure>
- Microsoft Azure 1.4 (April 2011)
 - Installer will ask for DBMS (MySQL or SQL Server)

Windows Azure Storage

- Scalable storage in the cloud
 - 100tb per storage account (3 replicas of data)
 - Auto-scale to meet massive volume and throughput
- Accessible via RESTful Web Service API
 - Access from Windows Azure Compute
 - Access from anywhere via internet
 - Supporting .NET Client Library
- Various storage types
 - Table - group of entities (name/value pairs)
 - Queue - Simple non-transactional message queue
 - Blob - Large binary storage
 - Drives - NTFS VHD mounted into Compute instance

SQL Azure Database

- Familiar SQL Server relational database model delivered as a service
 - Support for existing APIs & tools
 - Built for the cloud with high availability & fault tolerance
 - Easily provision and manage databases across multiple datacenters
- SQL Azure provides logical server
 - Gateway server that understands TDS protocol
 - Looks like SQL Server to TDS Client
 - Actual data stored on multiple backend data nodes
- Logical optimizations supported
 - Indexes, Query plans etc..
- Physical optimizations not supported
 - File Groups, Partitions etc...
- Transparently manages physical storage

Reporting & Data Sync

- SQL Azure Reporting

- SQL Server Reporting provided as a service
- Reports authored using existing tools (BIDS) and uploaded to the cloud
- Reports can have rich Data Visualizations (Maps, Charts, Tablix) and be exported to variety of rendering formats (Excel, Word, PDF)
- Reports can be rendered as part of an app using the Report Viewer control
- Directly view the reports in the browser
- Web Service interface to render and manage reports

- SQL Azure Data Sync

- Provides geo-replication
- Adds sync between SQL Server and SQL Azure
- Builds on Sync Framework

Summary

- **Windows Azure Platform** is a comprehensive **PaaS** offering including:
 - **Windows Azure**
 - **SQL Azure**
 - **Windows Azure AppFabric**
- Commercially available today in 41 countries and 6 data centers
- In the future Windows Azure & SQL Azure will be available on-premises as an appliance
- Continuing to expand the set of services and features

Windows Azure Platform Consumption Prices

Pay as you go and grow for only what you use when you use it



Windows Azure™



Microsoft®
SQL Azure™

**Elastic, scalable, secure, and highly available
automated service platform**

Compute

Per service hour

\$0.05-0.96/hour
+ Variable Instance Sizes

Storage

Per GB stored and transactions

\$0.15 GB/month
\$0.01/10k transactions

Web Edition

Per database/month

\$9.99/month
(up to 1 GB DB/month)

Business Edition

Per database/month

Starting at \$99.99/month
(10-50 GB DB/month)

Windows Azure AppFabric Service Bus and Access Control Service

Scalable, automated, highly available services for secure connectivity

Access Control

Per Message Operation

\$1.99/10k transactions

Service Bus

Per Message Operation

\$3.99/month per connection

MSR/XCG Cloud Computing Academic Research Engagements

Goals

Help to accelerate scientific exploration, discovery and results.

Broaden the research capabilities of scientists, foster collaborative research communities to accelerate scientific discovery globally.

Build partnerships with government-sponsored research agencies and university consortia to provide cloud services to academic and research communities worldwide

Help researchers interact with massively scalable data analysis tools and services directly accessible from their desktops, laptop or other mobile devices in the same way they now interact with Web search and other online resources.

Partnerships

- More than 70 projects under national programs with
 - NSF (32)
 - European Union (12 + 10)
 - Japan (6)
 - Australia (8)
 - China and Taiwan (. . .)

NSF (21 + 11)

- Inferring Pattern and Processes of **Genome Evolution** through Cloud Computing
- **GIS Vector Data Overlay Processing** on Azure Platform
- Porting the **Structure-Adaptive Materials Prediction** to the Azure Platform
- **Cooperative Developer Testing** with Test Intentions
- Towards automated and assurable **enterprise network migration**
- **Data Intensive Grid Computing** on Active Storage Clusters
- Moving Polarizable **Force Field Simulations** to the Microsoft Azure Platform
- Maximizing the Utility of Orthologs and Phylogenetic Profiles for Systems-Scale **Comparative Genomics**.
- Web-scale Language Modeling Features for **Machine Translation**
- Stork **Data Scheduler** for Azure
- Exploring **Social Classification** on Microsoft Azure
- Where the **Ocean Meets the Cloud**: Ad Hoc Longitudinal Analysis and Collaboration Over Massive Mesh Data
- Transforming Morphological Systematics From Desktop to Web Applications: Development of the Online Workspace **Morphobank.org** 3.0
- Semantic Web Informatics for Species in Space and Time

European Commission (12 + 10)

- Systems Biology
- Drug Discovery
- Bioinformatics
- Civil Engineering: Building Information Management.
- Civil Engineering: Structural Analysis of Buildings.
- Civil Protection and Emergencies.
- Data for Science: Aqua maps.
- UK: Project Horizon
- France: INRIA

Japan Info-plosion (6)

- Largest funding program from the Japanese Ministry of Education for informatics research on cloud computing in Japan.
- More than 300 attendees participated in the Info-Plosion symposium on 03/06/2011, in Tokyo, Japan.
- TSUBAKI, a text search engine developed by Prof. Kurohashi from Kyoto University, enables deep search by using NLP technology.
 - TSUBAKI **required more than 10,000 (CPU) cores on Windows Azure.**
 - Predicate-Argument Structure Analysis of Huge Web Corpora for Improving the Search Engine Infrastructure.
 - Indexing a large Web document collection with modality/factuality information document retrieval on the TSUBAKI search engine.
- Other projects:
 - A parallel workflow system for multi-cloud environment
 - Ubiquitous Content Management Technology R&D.
 - Inter-Cloud Large-Scale Data Transfer.



Microsoft®

© 2010 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation.

MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.